

¿ES CONSISTENTE LA HIPÓTESIS DEL CEREBRO EN UNA CUBETA?

Tobies Grimaltos Mascarós
(Universitat de València)

INTRODUCCIÓN

HACE algunos años escribí un artículo¹ en el que me oponía, entre otros, al argumento antiescético de Quine en «The Nature of Natural Knowledge»². Quine sostiene allí que:

En un mundo de *sense-data* inmediatos, sin cuerpos postulados..., una distinción entre realidad e ilusión no tendría lugar³.

Estimaba yo entonces que el argumento quineano compartía con otros argumentos contemporáneos la forma siguiente: Si algo X es concebible y sólo es concebible sobre la previa aceptación de Y, que es la condición de posibilidad de X, entonces si X puede ser el caso, Y tiene que serlo necesariamente. Por lo tanto, según el esquema de este argumento, lo que nos estaba diciendo Quine venía a ser que la existencia de ilusiones (X), es una prueba a favor y no en contra de la existencia de percepciones verídicas y, por tanto de objetos (Y), pues éstos son condición de posibilidad de aquéllas. Sólo en un mundo donde haya objetos puede haber ilusiones.

Consideraba yo entonces, pues, el argumento de Quine como un argumento de tipo transcendental, pero me oponía a él más o menos del siguiente modo. En mi opinión, este argumento daba un salto desde el concepto al objeto, desde

¹ «Cerebros en una cubeta y otros mitos escépticos», *VI.º Congrès de Filosofia del País Valencià*, Alicante, Institut «Juan Gil Albert», Ajuntament d'Elx, 1991; pp. 121-137.

² En Guttentplan, S. (ed.): *Mind and Language*, Oxford, Clarendon Press, 1975, pp. 67-81.

³ *Ibid.*, p. 67.

la posibilidad a la existencia, desde lo mental a lo físico. En general, me oponía a tal tipo de argumentos en tanto que pretendieran suponer una refutación del escepticismo, pues pensaba que lo único que conseguían mostrar era que, para que el concepto «ilusión» tenga sentido, ha de tenerlo el concepto de «objeto físico», y esto es algo que el escéptico puede aceptar al tiempo que niega que tal concepto tenga algún correlato real.

No estoy muy seguro de si mi opinión respecto del poder antiescético de este tipo de argumentos ha variado. Soy de la opinión de que el escepticismo es, en un sentido importante, irrefutable, invencible. Sin embargo, tampoco estoy seguro de que en aquella ocasión valorara adecuadamente la «fuerza transcendental» que tales argumentos (y en particular el de Quine) pueden tener, y esto es lo que me propongo abordar ahora.

1. EL CEREBRO EN UNA CUBETA

La versión de hipótesis escéptica más ampliamente debatida —casi habría que decir «más agobiadamente debatida»— en la literatura filosófica de nuestros días es la del cerebro en la cubeta. Ya se sabe, yo (cualquier yo que esto piense o lea) podría ser un cerebro (sin cuerpo) colocado en una cubeta llena de nutrientes que lo mantienen con vida y conectado mediante cables a una computadora enormemente potente que le suministra todas las experiencias sensoriales que tiene y que cree son del mundo real (del mundo en el que nosotros —yo— creemos vivir). Pero no hay tal mundo, fuera del cerebro mismo y su cubeta y la sala que comparte con la computadora (y, si se quiere, un genio, maligno o no, pero científico, que maneja todo el macabro entramado). No hay amigos, ni parientes, ni montañas, ni árboles, ni peces, ni dinero, ni empleos, ni siquiera de profesor de filosofía en alguna universidad.

La fuerza de esta hipótesis radica en que no haya ningún procedimiento al alcance del cerebro-víctima (que, recuérdese, podría ser quien esto lee) que le permita descubrir su error al creer que hay un tal mundo. Todo, por tanto, tiene que ser, para el cerebro engañado (desde su perspectiva), como les parece ser a quienes, si los hay, no son cerebros en una cubeta y viven en un mundo de objetos materiales. Pero, ¿puede esto ser realmente así? ¿Es consistente la hipótesis de los cerebros en la cubeta? Vale la pena que nos replanteemos la afirmación de Quine. Si no hay objetos físicos, ¿puede haber ilusiones?

Hemos dicho que para el cerebro en la cubeta todo tiene que parecer igual que si no lo fuera. El cerebro en la cubeta vive en una permanente ilusión, pero tiene que ser una ilusión cualitativamente indistinguible de la realidad. En la

realidad de la que nosotros nos suponemos habitantes también se sufren ilusiones y alucinaciones. Del mismo modo, aun dentro de una total ilusión, el cerebro en la cubeta debe poder sufrirlas. No todas sus «percepciones» deben ser internamente adecuadas, debe poder sufrir errores perceptivos dentro del permanente error. Debe haber ocasiones en las que sea internamente adecuado decir que percibe una silla y ocasiones en las que, a pesar de tener las experiencias pertinentes, no lo sea, pues todo en su vida experiencial debe ser idéntico (como tipo) a la vida experiencial de alguien que no es (o no fuera) un cerebro en una cubeta. En la vida del cerebro deben producirse experiencias que lleven a engaño y al fracaso en sus expectativas y pretendidas acciones. Habrá ocasiones (o debe poder haberlas) en las que lo que cree percibir no se adecua a sus subsiguientes percepciones: cree ver una silla y, al intentar sentarse, da con sus huesos en el suelo; tiene las experiencias correspondientes a dar con sus huesos en el suelo.

¿Existen criterios en el mundo del cerebro en la cubeta que permitan distinguir la percepción internamente verídica de la interna ilusión o alucinación? La respuesta puede parecer afirmativa. Y debe ser afirmativa si la hipótesis ha de poder formularse de un modo consistente. Pongámoslo ahora, pues, todo entre paréntesis y observemos si en el mundo posible del cerebro en la cubeta se dispone de los mismos criterios de que creemos disponer nosotros para poder decidir si algo es una percepción verídica o una alucinación. En primer lugar, parece que se dispone de los criterios fenomenalistas tales como la contigüidad y la semejanza de las impresiones sensibles. Si cree percibir una silla y los perceptos que del mismo lugar recibe a continuación no se parecen en nada a los anteriores, tendrá que concluir que ha sufrido algún tipo de ilusión. Por otro lado, están los demás, los otros, (o los datos sensibles que corresponden a los «demás»). Puede contar con el testimonio de los otros individuos. Si algo caracteriza perceptivamente a un objeto es la múltiple accesibilidad: un objeto es susceptible de ser percibido por varios sujetos al mismo tiempo. Si en condiciones normales de observación (buena luz, en un espacio abierto, etc.), los demás no consiguen tener experiencias semejantes, entonces es que probablemente no existe tal objeto.

Bueno, todo parece ser más o menos idéntico al mundo en el que existen objetos físicos. Pero, fijémonos, sólo hemos utilizado criterios compatibles con el fenomenalismo. Ahora bien, si sólo utilizamos criterios fenomenalistas, la hipótesis escéptica no tiene ninguna mordiente, pues los objetos existen en la medida en que son percibidos (o perceptibles) y no son más que conjuntos de percepciones (reales o posibles). Pero si esto es así, el mundo del cerebro en la cubeta es tan real como el nuestro. De acuerdo con los criterios fenomenalistas, seamos un cerebro en una cubeta o no, puede existir igualmente un mun-

do poblado de objetos físicos. Si la hipótesis escéptica supone algún problema es por la supuesta distinción que existe entre apariencia y realidad. Los objetos no son conjuntos de apariencias, son algo más. Sobre todo, son entidades que pueden existir aun cuando no son observadas ni observables. Lo que hay no tiene por qué corresponderse con lo que nosotros creemos que hay, por muchos que seamos los que creemos que eso es lo que hay. La hipótesis escéptica cobra sentido únicamente sobre la previa aceptación del realismo. Necesita asumir el realismo para poder tener algún sentido.

2. EL PRINCIPIO DE VERIFICACIÓN

Stroud acusa a ciertos argumentos transcendentales de presuponer el *principio de verificación* y de que sólo desde tal asunción pueden pasar desde las premisas psicológicas acerca de lo que creemos a la conclusión acerca de lo que hay⁴. Una curiosa coincidencia que merece que analicemos. Según Brueckner los argumentos transcendentales que Stroud explora adoptan estas dos formas⁵:

1) «Si atribuimos creencias a los hablantes (si creemos que tienen creencias con determinados contenidos), entonces también *debemos creer que hay* un mundo independiente de objetos perdurables con el que aquellos interactúan»⁶.

2) «Si los hablantes tienen creencias con contenidos determinados, entonces *hay* un mundo independiente de objetos perdurables con el que aquéllos interactúan»⁷.

El primer tipo de argumento supondría una forma de argumento transcendental moderado, mientras que la segunda sería la forma de un argumento transcendental más ambicioso. Stroud acepta (1) mientras que rechaza (2) y afirma que este último sólo es sostenible si aceptamos el principio de verificación, lo que supone renunciar al realismo. Así dice Stroud:

El principio de verificación en el que reposa el argumento es: si la noción de particulares objetivos tiene sentido para nosotros, entonces podemos saber en ocasiones que se cumplen ciertas condiciones, condiciones cuyo cumplimiento implica lógicamente que o bien los objetos continúan existiendo cuando no son percibidos o bien no lo hacen. El escéptico dice que nunca podemos justificar nuestra acep-

⁴ Stroud, B.: «Transcendental Arguments», en *The Journal of Philosophy*, vol. LXV, 9(1968), 241-256, y «Kantian Argument, Conceptual Capacities and Invulnerability» en Parrini, P. (ed.): *Kant and Contemporary Epistemology*, The Hague, Kluwer Academic Publisher, 1994, pp. 231-251.

⁵ Brueckner, A.: «Modest Transcendental Arguments», en *Philosophical Perspectives*, 10(1996), 265-280.

⁶ *Ibid*, p. 268. Cursiva nuestra.

⁷ *Ibid*, p. 269. Cursiva nuestra.

tación de la proposición de que los objetos continúan existiendo cuando no son percibidos... Si la afirmación del escéptico tiene sentido, debe ser falsa, puesto que si no pudiera saberse que la proposición es verdadera o saberse que es falsa, no tendría sentido. Esto se sigue del principio de verificación. Sin tal principio el argumento... no tendría fuerza; pero con este principio el escéptico es directamente y concluyentemente refutado, y no hay ninguna necesidad ulterior de proceder mediante un argumento indirecto o transcendental para exponer sus errores⁸.

Stroud afirma, pues, que sólo podemos refutar el escepticismo si abandonamos el realismo, esto es, si negamos la posibilidad de verdades que trasciendan la verificación. Abandonar esto supone, pues, aceptar que la existencia depende de nuestro poder cognitivo: una vez más, nada cambia por el hecho de que seamos cerebros en una cubeta o no.

Nos encontramos ante un curioso fenómeno: la hipótesis escéptica sólo puede formularse sobre la previa aceptación del realismo y sólo puede ser vencida si lo abandonamos. El realismo juega, por tanto, la baza definitiva en esta cuestión. Pero esto nos debe dar que pensar. Stroud está acusando a cierto tipo de argumentos transcendentales de vencer al escepticismo sólo a través del movimiento ilegítimo de integrarse en el marco de su hipótesis, de adoptar como criterios de realidad los únicos criterios que tal marco permitiría. Pero a su vez el escepticismo está promulgando que los criterios de realidad no pueden coincidir con los que están disponibles a la víctima de su experimento mental (cualquiera de nosotros o todos nosotros tomados de uno en uno); si no, su hipótesis no tendría ninguna fuerza. Así, la estrategia del escéptico viene a ser la siguiente:

(a) Acepta los criterios realistas como criterios de existencia.

(b) Intenta mostrar mediante una hipótesis que no es posible demostrar que algo cumple tales criterios.

Por su parte, la estrategia de los argumentos transcendentales del tipo de (2) sería la de abandonar tales criterios, con lo que cualquier argumento se hace inútil.

En opinión de Stroud, los argumentos transcendentales moderados son viables, pero no dan el salto de lo psicológico a lo real, desde el concepto al objeto o desde la creencia a su verdad; los argumentos transcendentales más ambiciosos, los que pretenden pasar de lo psicológico a lo real, fracasan. El argumento

⁸ Cf. nota 4.

de Quine parece ser, en ese respecto, del segundo tipo; recordemos mi acusación inicial. Sin embargo, lo que me propongo mostrar ahora es que algún tipo de argumentación que siga las líneas del de Quine puede servir para mostrar cierta inconsistencia del planteamiento escéptico. ¿La posibilidad de la ilusión necesita como condición necesaria la existencia de objetos físicos? Creo que la hipótesis del cerebro en la cubeta necesita la previa aceptación del realismo para poder ser formulada, necesita de un criterio de verdad y existencia independiente de nuestras capacidades cognitivas y a la vez lo hace imposible, al menos en algunas ocasiones. La situación que describe la hipótesis es imposible desde los planteamientos realistas: no cumple con sus criterios.

3. CASOS EN LOS QUE LA ILUSIÓN ES IMPOSIBLE

Imaginemos el siguiente caso. Un alemán pasa unos días en Oxford. Visita el Merton College. Solo, sin la compañía de nadie, explora un pasillo desierto, abre una puerta, echa una mirada y, durante dos escasos segundos, observa una mesa en el fondo de la habitación. Al cabo de muchos años le da por preguntarse si vio realmente aquella mesa o si se trató de una alucinación. Ahora, claro está, no dispone de otras experiencias con las que cotejar aquélla. No puede volver a la habitación e intentar observar otras partes de la mesa o tocarla. No hubo testigos, ni hay nadie a quien preguntar. Supongamos que aunque volviera a Oxford no vive ya nadie que pueda confirmarle o no que el 24 de septiembre de 1926, a las 11 de la mañana, había o no una mesa en aquella habitación. No hay ningún criterio, de los que pueda servirse un fenomenalista para determinar que la percepción era verídica o no, a su disposición. ¿Pudo a pesar de todo tratarse de una alucinación? ¿Tiene sentido la posibilidad de que se tratara de una alucinación? Claro que sí. ¿De qué depende que se tratara o no de una alucinación? Depende de que existiera realmente una mesa allí con la que tuviera un contacto perceptivo adecuado; depende de que existiera una mesa que causara del modo adecuado su experiencia perceptiva, de que existiera una mesa con la que estableció un vínculo informacional apropiado..., o de que, en cambio, la causa de su experiencia no fuera la mesa sino algún desajuste interno de su sistema receptor. Aunque no haya manera de verificar si fue o no una alucinación, la posibilidad de que fuera una alucinación tiene sentido. Yo, como narrador omnisciente, podría decir ahora que se trató efectivamente de una alucinación.

Imaginemos ahora que tal caso no le ocurrió a un alemán, sino que es un cerebro en una cubeta quien cree haber vivido tal experiencia y quien se pregunta ahora si fue una alucinación. ¿Pudo ser una alucinación? Sinceramente creo que tal posibilidad no tiene sentido. Ahora no existen criterios que permi-

tan dirimir la cuestión y, si no hay criterios, la cuestión no tiene sentido. Donde no hay objetos físicos, hay al menos algunos casos en los que la ilusión no es posible. Imaginemos que el cerebro tiene las experiencias correspondientes a las de volver al Merton College. Supongamos que cuando ahora tiene las experiencias correspondientes a volver a abrir la puerta de aquella habitación, no tiene las experiencias correspondientes a ver aquella mesa o, mejor, tiene las experiencias que a cualquiera nos harían decir que no hay ninguna mesa. Si hemos dicho antes que no hay nadie que pueda confirmarle o no que el 24 de septiembre de 1926, a las 11 de la mañana, había o no una mesa en aquella habitación, ello significa, en el caso del cerebro en una cubeta, que jamás lo hubo. Si ahora no hay testigos, nunca los hubo, pues el hecho de que el cerebro en la cubeta no entrara en contacto con ellos supone que nunca existieron, que no forman parte del programa de la computadora que crea su mundo. Ni siquiera yo, como narrador omnisciente, puedo ayudar.

Quizá veamos mejor lo que pretendo decir si ampliamos nuestro enfoque a los casos de ilusión. Porque, es cierto, esto puede ser ampliado a casos de ilusión. Supongamos, que en aquella ocasión, la mesa le pareció verde. Imaginemos además que su recuerdo es correcto: efectivamente, la mesa le pareció verde entonces. Pero ahora se pregunta si realmente la mesa era verde o aquella apariencia era debida a que la luz que alumbraba la habitación no era la habitual. Se pregunta si quizás los cristales de las ventanas no serían verdes y por eso la mesa presentaba aquel color. Pero él no miró hacia los cristales, no tuvo perceptos de los cristales. Si el individuo de nuestro ejemplo es un cerebro en una cubeta, una vez más, la posibilidad de que fuera una ilusión no tiene sentido. Si la computadora no le suministró datos sensoriales de las ventanas, la posibilidad de que la mesa fuera verde a causa de los cristales de tales ventanas no tiene sentido. Si no hay más mesa ni más cristales que los que él creyó ver (o no creyó ver), la «mesa» existió mientras él la vio y era del color que él la vio, pues no había más mesa ni más propiedades de la mesa que las que él vio. Ni siquiera un narrador omnisciente tiene criterios para decidir si era o no una ilusión, porque, simplemente, ya no hay distinción posible. Un narrador omnisciente no puede tener (en este contexto), más datos que el programa de la computadora, y en el programa de la computadora (bajo el supuesto de que su única razón de ser es la de presentarle al cerebro una representación del mundo) no hay datos sobre las ventanas y sus cristales. Supongamos que nuestro individuo decide volver y aquella habitación ha sido reformada. Sólo existió tal como él la vio cuando él la vio y constaba de los elementos que él vio.

4. ACLARACIONES Y OBJECIONES

Antes de seguir, me gustaría aclarar que, con esto, no pretendo refutar al escepticismo. Mi argumento podría mostrar ciertos problemas en ciertas versiones del argumento del cerebro en la cubeta, pero no es ningún obstáculo para otras. No sirve en contra de una situación como la que narra John Pollock en *Contemporary Theories of Knowledge*⁹, en la que a un ser humano que ha vivido como tal durante años se le extrae el cerebro y se lo coloca en una cubeta. Creo, que a pesar de todo, la posibilidad de que yo (cualquiera de nosotros) sea un cerebro en una cubeta sigue siendo una posibilidad epistémica¹⁰. Que no haya posibilidad de ilusión interna en algunos casos no significa que todo no sea más que una ilusión. Quizá lo que nosotros suponemos la realidad no sea sino el producto de los *outputs* de una computadora conectada a nuestro cerebro mediante electrodos. Tal vez el planteamiento escéptico de la hipótesis del cerebro en una cubeta sea inconsistente (necesita adoptar el realismo mientras su naturaleza misma lo niega), pero sigue siendo una posibilidad epistémica. Quizá, después de todo, no haya sustancias externas que causen nuestra percepción de la manera que nosotros pensamos que lo hacen y que sólo haya unas ilusiones más perdurables que otras. Dicho de otro modo, que la hipótesis del cerebro en la cubeta haga imposible la existencia de ciertas ilusiones, porque la ilusión coincidiría con la realidad, no supone que la realidad sea otra cosa que la creación, por parte de una computadora, de una serie de experiencias perceptivas en un cerebro sin cuerpo sumergido en una cubeta.

De todos modos y dicho lo anterior, incluso se pueden poner objeciones a la afirmación más débil, la que sostiene que según los criterios fenomenalistas no hay posibilidad de determinar si en los casos descritos se trató o no de una ilusión o una alucinación. Si existe algún criterio para dilucidar esta cuestión, si la pregunta tiene una respuesta con independencia de que se pueda conocer o no, entonces la hipótesis es perfectamente congruente con el realismo. Y esa respuesta, pueden decir muchos, existe. Pues, se nos puede decir, incluso desde los criterios fenomenalistas existe un criterio para decidir si se trata o no de ilusiones. Alguien nos puede acusar de haber olvidado un aspecto crucial del fenomenalismo: los condicionales subjuntivos que determinan la existencia de los objetos. La máxima por excelencia del fenomenalismo es que los enunciados que afirman la existencia de los objetos físicos son equivalentes en su significado a los enunciados que describen sensaciones posibles. Como dice Richard Fumerton en la *Routledge Encyclopedia of Philosophy*¹¹, de acuerdo con el fenomenalismo

⁹ Londres, Rowman & Littlefield, 1986, pp. 1-3.

¹⁰ Permítaseme decir, de pasada, que hay ciertas posibilidades epistémicas que son metafísicamente imposibles, incluso pueden ser lógicamente imposibles.

¹¹ Londres y Nueva York, Routledge, 1998. Cf. la entrada *Phenomenalism*.

lismo, «debemos tomar los enunciados sobre objetos físicos como enunciados que hacen aserciones sobre las sensaciones que alguien *tendría si* esa persona tuviera ciertas otras sensaciones». Por tanto, existiría un modo de determinar si el sujeto de nuestro experimento mental sufrió o no, en cada caso, una delusión. Si hubiera mirado en tal o tal otra dirección, ¿hubiera tenido tales y tales otras sensaciones? Y esta pregunta tiene una respuesta, piensa nuestro objetante, con independencia de que la conozcamos o no. La imposibilidad de determinar si se trató o no de una delusión es una imposibilidad epistémica, y no, tal como yo he insinuado, una imposibilidad metafísica.

Esta objeción me parece poderosa y hará que, luego, hayamos de hacer algunas precisiones añadidas. Pero, antes, hemos de tener en cuenta algunas cosas. Me parece que en cierto modo la posible verosimilitud de esta estrategia fenomenalista depende de la postulación de objetos externos al sujeto (y a otros sujetos). Depende de la existencia de unas «posibilidades permanentes de sensaciones», como dijo J. S. Mill¹², que no sólo son intersubjetivas, sino que de algún modo han de ser objetivas. Los condicionales contrafácticos del fenomenalismo parecen tener una respuesta en la medida en la que hay un mundo de entidades permanentes que les dan una respuesta. En el mundo del fenomenalista, a pesar de todo, la pregunta sobre el color de la mesa o sobre si la mesa era real, depende de unos hechos del mundo que garantizan unas experiencias de un sujeto en el caso de que el sujeto hubiera realizado las acciones adecuadas: en la medida en que hubiera cumplido con las condiciones especificadas en el antecedente del condicional.

Pero el caso del cerebro en la cubeta, en los ejemplos que hemos descrito, se parece más al capítulo de una novela que al mundo real. Para ver qué es lo que quiero decir, imaginemos la siguiente situación:

Juan baja por las escaleras. Lleva unos zapatos de piel. María lo mira.

En el mundo real, la pregunta «¿De qué color eran los zapatos de piel de Juan?» tiene una respuesta, con independencia de que ni María ni nadie los haya visto y la sepa responder. El condicional contrafáctico «Si María hubiera mirado en dirección a los zapatos de Juan hubiera tenido una sensación de marrón» podría ser verdadero (o falso: tiene un valor de verdad determinado)¹³. El

¹² *An Examination of Sir William Hamilton's Philosophy*, Green, Longmans, 1889.

¹³ Estoy dejando ahora a un lado posibles excepciones que tienen que ver con la no satisfacción de la cláusula *ceteris paribus* o de la cláusula «en condiciones normales de observación» por no hacer más farragosa la exposición. Así no tengo en cuenta objeciones como la de Chisholm de que la existencia de un objeto físico no implica la existencia de dato sensorial alguno. Y no lo tengo en cuenta, porque lo que digo arriba es una concesión al fenomenalista. Aunque el condicional subjuntivo no tuviera un valor de verdad, debido a objeciones como ésta, la pregunta «¿De qué color son los zapatos de Juan?» sí que la tendría y esto es todo lo que importa; el resto son concesiones al rival.

mundo hace verdadero (o falso) ese condicional a pesar de la ignorancia (o no) de María o, incluso, de todos. Pero imaginemos que la situación corresponde a un pasaje del capítulo de una novela en la que el autor sólo dice: «Juan baja por las escaleras. Lleva unos zapatos de piel. María lo mira». ¿De qué color son ahora los zapatos? Para que la pregunta anterior tenga una respuesta, hay que volver a escribir el capítulo añadiendo ese dato. Ésta es la situación más similar a la del cerebro en la cubeta. Para que el contrafáctico tenga un valor de verdad, hay que crear un nuevo mundo o, por lo menos, completarlo. De otro modo, en el mundo del cerebro en la cubeta no existe lo que los anglosajones llaman «*truth-maker*» para tal contrafáctico.

Ahora bien, si esto es verdad, entonces la cantidad de creencias y afirmaciones del cerebro a las que les sucedería lo mismo crece de un modo desmesurado. ¿Hay un número determinado de evidencias que implique la verdad de una proposición sobre un objeto o hecho físico? Y si lo hay, ¿cuántas veces las conseguimos? En un mundo de seres humanos con cuerpo y objetos físicos, es posible obtener un considerable número de evidencias que avalen la creencia de que hay una mesa verde ante mí (o un teléfono blanco, o un cuervo negro) y descubrir con posterioridad que uno estaba equivocado; incluso es posible estar equivocado y no descubrirlo jamás: por ejemplo, porque se cambió para siempre de contexto y no se tuvo ocasión de que nuevas evidencias nos mostraran el error. En el mundo «real», de todos modos, los contrafácticos sobre posibles experiencias posteriores tienen siempre un valor de verdad¹⁴, puesto que la pregunta sobre si se trata o no de una mesa verde, un teléfono blanco o un cuervo negro, tienen siempre una respuesta, que a veces es negativa.

En el mundo del cerebro en la cubeta, cuando hay un número de evidencias que no *implicarían internamente* la existencia de la mesa o del teléfono, pero tampoco ha habido experiencias subsiguientes que ratifiquen o desmientan las anteriores, ¿tendría sentido la posibilidad de que no existiera la mesa o el teléfono? ¿Qué valor de verdad tendría el contrafáctico relativo a las experiencias no tenidas? Y si ese contrafáctico no tiene valor de verdad, ¿ello no implica que la aserción de la existencia de la mesa (o del cuervo o del teléfono) en tales circunstancias tampoco la tiene? Parece, pues que en el mundo (o quizás sea mejor decir no-mundo) del cerebro en la cubeta, existir implica ser percibido, incluso si la inversa no se da. Imaginemos que en ese mundo del cerebro en la cubeta puesto entre paréntesis se produce la siguiente situación. La víctima del experimento va paseando por un bosque y ve un árbol caído. Sigue su camino hasta casa. Cuándo llega se pregunta si aquel árbol fue talado o si más bien fue

¹⁴ He dicho que ignoraba ciertos problemas que podrían hacer falsa esta afirmación.

derribado por el viento o por un rayo. Pero él no prestó atención (no miró) al lugar por donde el árbol estaba separado del suelo, así que no vio marcas de sierra o de quemaduras, o las astillas que se producen cuando el viento quiebra los árboles. Se promete que tan pronto pueda volverá al bosque para comprobarlo. Cumple su promesa al cabo de una semana, pero el árbol ya no está: alguien lo ha retirado. Su pregunta, acerca de si el árbol fue talado o no, ¿tiene una respuesta?

Justamente ésta es la línea que puede seguir nuestro objetante. Puede objetar a todo esto que nuestro argumento reposa en una determinada concepción de la computadora a la que está conectada el cerebro (o de quienes la controlan). Mi argumento reposa en el supuesto de que la computadora no tenga un programa completo. El programa no tiene datos sobre ciertas posibilidades que no se han convertido en efectivas. Así, no tiene más datos sobre la mesa y la «vida» de la mesa o sobre los cristales de la sala porque, entre otras cosas, no ha habido necesidad de aportarlos, dado que el cerebro no ha tomado la decisión de entrar en la habitación o mirar los cristales no ha obtenido más experiencias relevantes y, por consiguiente, la computadora o sus controladores no han tenido que crearlos. Pero nuestro objetante puede decir que ello no tiene por qué ser así. La computadora podría tener un programa completo con todas las experiencias que el cerebro podría obtener en todas las circunstancias posibles.

Ante esto, hay dos posibles movimientos. El primero consiste en decir que no hay ninguna necesidad de que la computadora y su programa sean como mi oponente dice. De hecho, parece que la mayoría de los planteamientos de la hipótesis del cerebro en la cubeta cobran verosimilitud del supuesto implícito de que el programa actúa de acuerdo con el *feedback*, sin tenerlo todo preterminado. Tendemos a pensar que la computadora no contiene *a priori* en su programa «todos los mundos posibles» que las distintas posibles decisiones del cerebro pudieran originar. Y eso, creo, tiene una razón de ser que tiene que ver con el segundo movimiento que podemos hacer. El segundo movimiento, quizá más interesante, consiste en aceptar que una hipótesis escéptica basada en este otro tipo de posible computadora sería, aparentemente al menos, más poderosa. En ese caso, parece que todos los contrafácticos tendrían respuesta, tendrían un valor de verdad. Pero entonces la hipótesis del cerebro en la cubeta, se convierte más bien en la hipótesis de un genio maligno omnipotente. Un genio maligno que ha creado todas las posibilidades (si ello tiene sentido) desde el principio. Cada posible escenario está prediseñado y listo para dar satisfacción a cualquiera de las libres decisiones del cerebro en la cubeta. En tal caso sí que sería posible distinguir entre internas ilusiones e internas realidades. Ahora bien, las creencias que no fueran internamente falsas, es decir, las creencias que se adecuaban a las apariencias, ¿en qué sentido serían falsas? En el sentido, se dirá,

de que no se adecuan a la realidad externa. Pero fijémonos que si el programa es completo, esta distinción se va diluyendo de un modo cada vez mayor. Ahora bien, ¿tiene sentido un programa completo que no sea él mismo la realidad? ¿Sería posible algo semejante?

Volvamos a nuestro ejemplo del árbol. ¿Tiene sentido que la computadora haya creado toda una vida del árbol desde que era una semilla hasta que fue un árbol adulto, hermoso y grande, al margen completamente de la conciencia del cerebro en la cubeta? Dejando ahora a un lado cuestiones de economía computacional, ¿qué podría significar esto? ¿Puede un programa de computadora tener respuesta para todas las preguntas? ¿Tiene sentido que un genio maligno cree toda una realidad virtual y al margen de la conciencia del individuo, sin que ello no sea «la realidad»? Pero me da la sensación de que incluso en el caso del programa completo muchas afirmaciones seguirían sin tener un valor de verdad interno. ¿Cuándo cayó el árbol de nuestro último ejemplo? ¿Lo vio alguien caer? ¿De cuántos decibelios era el sonido que produjo al chocar contra el suelo? ¿Cuántas hojas tenía? ¿Sufrió alguna plaga alguna vez?, ¿de qué? ¿Cuántos insectos constituían la plaga? Cualquiera de las posibles respuestas a cualquiera de estas preguntas, ¿tendría un valor de verdad? Y si lo tuvieran, ¿tiene sentido una «realidad» tan completa y paralela a la vida mental del cerebro? Si ha de ser posible la distinción interna entre ilusión y realidad, ¿cómo ha de ser el «mundo» del cerebro en la cubeta? ¿Es realmente posible algo así?¹⁵

¹⁵ La investigación conducente a este trabajo ha sido financiada por la DGES como parte de los proyectos PB96-0764 y PB96-1091-C03-02. Agradezco la ayuda y el estímulo prestados por esta institución. Diversos compañeros y amigos han hecho valiosos comentarios, críticas y sugerencias a versiones anteriores de este trabajo; son especialmente de agradecer los que hicieron: Josep Lluís Blasco, Josep Corbí, Christopher Hookway, Carlos Moya, Jordi Valor y José Zalabardo.